

## Análisis clúster: un primer paso sobre técnicas cuantitativas en Arqueología de la Arquitectura \*

### *Cluster analysis: a first step in quantitative techniques in Archaeology of Architecture*

Agustín Azkarate<sup>1</sup>, Ismael García-Gómez<sup>2</sup>, Amaia Mesanza-Moraza<sup>3</sup>  
Miembros de GPAC, Grupo de Investigación en Patrimonio Construido de la Universidad del País Vasco (UPV/EHU)

#### RESUMEN

La propuesta que se hace en este artículo parte del convencimiento de que existen ciertos métodos de análisis en arqueología que podrían mejorar notablemente si incorporaran las técnicas cuantitativas; uno de esos métodos es sin duda el de la lectura estratigráfica de alzados. A lo largo de estas líneas expondremos cual es nuestro bagaje al respecto, haciendo un breve recorrido que, si bien partirá de los primeros experimentos más intuitivos, se centrará prioritariamente en nuestros últimos ensayos de carácter matemático-estadístico. En el texto se apreciará cómo estamos experimentando con métodos de captura masiva de información geométrica que después, mediante programación, sometemos a una minería de datos basada en el empleo de algoritmos propios de las técnicas de análisis multivariante. Aportamos finalmente nuestra reflexión sobre un futuro en el que prevemos que la lectura estratigráfica de alzados alcanzará un grado de automatización muy próximo a los sistemas expertos y la inteligencia artificial.

**Palabras clave:** análisis estratigráfico; registro topográfico; arqueología cuantitativa; programación; estadística; análisis multivariante; automatización; sistemas expertos.

#### ABSTRACT

The proposal made in this article is based on the conviction that there are certain methods of analysis in archaeology that could be significantly improved if they incorporated quantitative techniques; one of those methods is undoubtedly that of the stratigraphic reading of elevations. In the course of this document, our background in this regard will be explained, by means of a brief summary. Although the starting point will be based on the initial more intuitive experiments, it will focus primarily on our latest mathematical-statistical trials. The text will identify how we are experimenting with methods of massive capture of geometric information, which through programming is later subjected to data mining, based on the use of multivariate analysis techniques with proprietary algorithms. Finally, we reflect on the future in which we envisage that the stratigraphic reading of elevations will reach a degree of automation very close to expert systems and artificial intelligence.

**Keywords:** stratigraphic analysis; topographic record; quantitative archaeology; programming; statistics; multivariate analysis; automation; expert systems.

Recibido: 02-03-2018. Aceptado: 05-04-2018. Publicado online: 20-12-2018

#### Cómo citar este artículo / Citation

Azkarate, A., García-Gómez, I. y Mesanza-Moraza, A. 2018: "Análisis clúster: un primer paso sobre técnicas cuantitativas en Arqueología de la Arquitectura", *Arqueología de la Arquitectura*, 15: e066. <https://doi.org/10.3989/arq.arqt.2018.014>

**Copyright:** © 2018 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

\* Trabajo realizado en el marco del Proyecto de investigación "La Arqueología de la Arquitectura: nuevos retos metodológicos y su aplicación en Europa y América Latina" (HAR2015-64439-P), financiado por el Ministerio de Economía y Competitividad y Fondo Europeo de Desarrollo Regional (FEDER).

<sup>1</sup> [agustin.azkarate@ehu.eus](mailto:agustin.azkarate@ehu.eus) / ORCID iD: <http://orcid.org/0000-0002-0240-4960>

<sup>2</sup> [ismael.garcia@ehu.eus](mailto:ismael.garcia@ehu.eus) / ORCID iD: <http://orcid.org/0000-0002-1464-8268>

<sup>3</sup> [amaia.mesanza@ehu.eus](mailto:amaia.mesanza@ehu.eus) / ORCID iD: <http://orcid.org/0000-0001-5162-8955>

## INTRODUCCIÓN

Durante muchos años, la Arqueología de la Arquitectura (AA) tuvo que medir sus fuerzas con una tradición historiográfica, dominante en el pasado siglo, en la que la analogía o el estilo constituían el principal de los argumentos para adscribir una obra arquitectónica a uno u otro periodo histórico. En aquel contexto, era natural que el rigor en la aplicación de los principios estratigráficos en arquitectura se convirtiera en una prioridad ineludible (Azkarate 2013: 280). Pero ocurrió —como ocurre en otras tantas disciplinas cuando están todavía en proceso de maduración— que la necesidad de explicar la nueva metodología de análisis, la urgencia por describir el *cómo* de algo que, en ciertos casos, estaba siendo practicado de un modo voluntarioso y probablemente improvisado, hizo olvidar el *por qué* y el *para qué* de la renovación que se estaba propugnando.

Quizá debido a ello, estos últimos años estamos notando un cierto estancamiento creativo. Tenemos a veces la sensación de que se ha instalado imperceptiblemente una ortodoxia que empuja al anquilosamiento a quienes trabajamos en el ámbito de la AA. Son dos, en opinión nuestra, las razones que pueden estar coadyuvando a esta situación:

*La equiparación de la AA con la lectura estratigráfica de paramentos.* La capacidad del método estratigráfico para decodificar y ordenar diacrónicamente la complejidad constructiva de un edificio histórico fue advertida pronto, estableciéndose una temprana relación con el ámbito de la restauración (*ibidem*: 282). Ello explica probablemente el éxito absoluto de la “*lettura stratigrafica muraria*”. La fácil legibilidad de sus resultados, su inmediata aplicabilidad a los estudios previos exigidos en los trabajos de restauración y rehabilitación, la existencia de unos protocolos cada vez más normativizados, la difusión de estos a través de los estudios de grado y postgrado, etc. han conducido indefectiblemente a su difusión.

La segunda de las razones del estancamiento de la AA tiene que ver, en opinión nuestra, con la *marcada preferencia de muchos de los especialistas en AA por la arquitectura de Medievo* (mucho más recientemente también por la arqueología clásica). Ello se debe, sin duda, a la influencia de los primeros maestros —fueran italianos o españoles— que, siendo impulsores de la AA, eran simultáneamente medievalistas reconocidos. Es muy probable que, en esta marcada especialización, haya influido también la temprana vinculación de la AA con el ámbito de la restauración arquitectónica, centrada frecuentemente en arquitecturas de origen medieval (*ibidem*: 292-294).

En este contexto, las consecuencias del trinomio reduccionista que parece imperar en Europa (AA=*stratigrafia muraria*=arquitectura medieval) son evidentes. Creemos que todo ello ha podido empobrecer el bagaje instrumental de una disciplina que nació con una vocación fundamentalmente integradora, representada por un amplio abanico de estudios arqueológicos que T. Mannoni unificó bajo el paraguas común de la *Archeologia globale* y que quedan oscurecidos por el monocultivo metodológico al que nos hemos referido.

Es precisamente sobre esta depauperación sobre lo que queremos reflexionar en este trabajo, aunque nos centraremos en esta ocasión en un aspecto muy específico, planteándonos críticamente si en los últimos decenios —subyugados quizá por la llamada “revolución estratigráfica” de los años 70 y 80—, no hemos relegado o minusvalorado otras herramientas (tipológicas y cuantitativas) que a la postre van a resultar, probablemente, tanto o más imprescindibles para el futuro de nuestra disciplina que la propia estratigrafía. Esta es —expresada de manera quizá un tanto brusca— la idea que defenderemos en las líneas que siguen.

## TIPOLOGÍA Y ESTRATIGRAFÍA EN ARQUEOLOGÍA DE LA ARQUITECTURA

No es la primera vez que desde GPAC reflexionamos sobre estas cuestiones. El hecho de haber tenido que enfrentarnos en fechas tempranas (1996-1998) al edificio histórico probablemente de mayor envergadura hasta ahora estudiado con la metodología propia de la AA (Caballero 2010: 118) nos condujo a encarar problemas de extraordinaria complejidad que ni siquiera habíamos previsto, acostumbrados como estábamos —a inicios de la década de 1990— a protocolos que habían mostrado su operatividad en elementos patrimoniales mucho más sencillos (Azkarate *et al.* 1995). El volumen de un edificio como la Catedral Vieja de Vitoria-Gasteiz y, sobre todo, su enorme complejidad constructiva, acabaron convirtiéndose en factores que nos empujaron a la innovación y la experimentación con diversas estrategias de estudio tanto “interfaciales” como “tipológicas”.

Simplificando al máximo podríamos decir que las primeras tienen como objeto definir la *envolvente* —o *interfaz*— *tridimensional del estrato*, una superficie envolvente más conceptual que real.

Las segundas, por el contrario, buscan la individualización del estrato a partir de la identificación de

los rasgos internos que caracterizan su contenido, no los de su envolvente. Aunque normalmente no seamos conscientes de ello, esta técnica es la que más empleamos los arqueólogos. Basta ver los epígrafes que componen nuestras fichas de registro para comprobar cómo, cuando llega el momento de traducir la realidad observada al lenguaje alfanumérico, el investigador recurre a la valoración, medición o consideración por separado de un cierto número de variables que son las que, conjuntamente, se consideran definitorias, no sólo del estrato que se está describiendo, sino de la globalidad de hechos estratigráficos de su mismo género. Por ejemplo, cuando se describe un paquete estratigráfico se ponderan variables como (1) “compactación”, (2) “color”, (3) “composición”, “tamaño de partícula”, (4) “intrusiones”, (5) “dimensiones”, etc., mientras que, cuando se habla de un elemento constructivo, se valoran características como (1) “materiales”, (2) “dimensiones”, (3) “talla”, (4) “aparejo” o (5) “mortero”, entre otras. La estrategia tipológica *no se preocupa tanto por las fronteras del estrato cuanto por su sustancia esencial*, por esa materialidad particular que nos permite diferenciar una realidad de otra caracterizada por unos rasgos notablemente diferentes<sup>4</sup>.

Sea como fuere, la “estrategia interfacial” no está puesta en cuestión y queremos que ello quede meridianamente claro para evitar equívocos en el futuro. Su eficacia a la hora de individualizar estratos queda fuera de toda duda; pero ello no impide observar ciertas limitaciones en las que sin embargo la “estrategia tipológica” da notables muestras de ser más resolutive. Hablamos fundamentalmente de contextos construidos de gran complejidad estratigráfica —como pueda ser el de una catedral o un casco urbano—, contextos que además de estar compuestos por miles de estratos, ocupan una superficie que es tendencialmente más extensa que profunda (o si se prefiere más tendente a la horizontal que a la vertical), factor que reduce la eficacia informativa de la interfaz con respecto a la caracterización estratigráfica por vía tipológica (debido en gran medida a que esta última permite establecer equivalencias entre estratos que no tienen una relación física directa). Fue la observación de estas limitaciones la que hace años nos animó a empezar a trabajar en una propuesta de método que explotase las potencialidades de la “estrategia tipológica”.

<sup>4</sup> Véanse las importantes reflexiones de L. Caballero a este respecto (2009 y 2010).

## CLÚSTERES Y “CLUSTERINGS”

El proyecto de la catedral de Santa María de Vitoria-Gasteiz vino a demostrar cómo dos sistemas de trabajo —en apariencia diferentes y considerados frecuentemente como antagónicos e incluso antitéticos— se complementaban entre sí (Caballero 2009: 12). El primero de ellos —*de naturaleza tipológica*—, operaba decodificando (deconstruyendo) las partes constitutivas de un edificio mediante la individualización de sus distintas variables para observar luego cómo se asocian entre sí hasta conformar conjuntos constructivos (“clústeres”<sup>5</sup>) de naturaleza sincrónica. El segundo —*de carácter estratigráfico*— analizaba aquellos clústeres que tuvieran relación física entre sí para descubrir, a través de sus interfaces, su ubicación en la secuencia constructiva del edificio; su naturaleza era, por tanto, diacrónica, aunque la secuencia cronológica fuera de momento “relativa”. Sólo tras la aplicación de otras herramientas de análisis (fuentes escritas sobre soportes diversos, numismáticas, análisis arqueométricos, etc.) se alcanzará la cronología absoluta deseable, que una vez alcanzada, permitirá la creación de cronotipologías de alcance local y regional<sup>6</sup>.

Los resultados de esta nueva propuesta fueron tan satisfactorios que pudieron ser testados con éxito en distintos contextos de aplicación, algunos de notable complejidad, como la arquitectura religiosa medieval del territorio alavés que analizó casi 400 iglesias (Sánchez Zufiaurre 2007) o el gran conjunto preindustrial de Salinas de Añana, con 5648 eras de sal, 2040 muros, 848 pozos, 258 almacenes y varios kilómetros de canales (Plata 2007).

En todos ellos aparece con frecuencia el término “clúster”. Con este anglicismo, tan difundido en la literatura científica, queremos expresar en castellano la idea de conjunto, agrupamiento, racimo o asociación. A partir de esta idea de “grupo” como denominador común, el anglicismo puede adquirir significados específicos según se aplique en la informática, la industria o las ciencias, por citar solo algunos ejemplos. ¿Cuál es el sentido que

<sup>5</sup> Adoptamos el sustantivo “clúster”, con tilde y plural “clústeres”, recogido con grafía hispanizada en el *Diccionario del español actual*, de Manuel Seco, Gabino Ramos y Andrés Olimpia (Madrid: Aguilar, 1999, 2 vol.).

<sup>6</sup> La primera vez que esta metodología se presentó en público fue en el *I Congreso Europeo de Restauración de Catedrales Góticas*, celebrado en Vitoria-Gasteiz en 1998 (Azkarate y Lasagabaster 2001); de nuevo en el *V Congreso de Arqueología Medieval Española*, celebrado en Valladolid en 1999 (Azkarate 2001); y finalmente en *I Seminario Internacional de Arqueología de la Arquitectura* llevado a cabo en Vitoria-Gasteiz en 2002 (Azkarate 2002); para una formulación más reciente puede consultarse Azkarate 2010.

otorgamos nosotros al concepto? De la misma manera que en astronomía se recurre al concepto “clúster” para referirse a un grupo de estrellas vecinas, unidas por gravitación mutua, que tienen esencialmente la misma edad y composición y, por lo tanto, supuestamente un origen común, en nuestros primeros trabajos recurrimos al concepto de “clúster constructivo” para referirnos a un grupo de “variables” (tanto técnicas como formales), unidas en su materialidad por la naturaleza construida de una fábrica, a las que cabe presuponer también la misma edad y, por lo tanto, un origen común, de acuerdo con el principio de la homogeneidad formal que todo acto constructivo coetánea conlleva. Como pudimos comprobar en la experiencia de la catedral de Santa María, la identificación (1º) de las variables técnico-formales, su georreferenciación (2º) en el espacio y el descubrimiento (3º) de la combinación tridimensional de estas variables entre sí, puso ante nuestros ojos —casi como en un ejercicio de prestidigitación— la existencia de distintas fases constructivas que, articulándose entre sí, componían de manera coral la torturada historia de un templo catedralicio.

Sin embargo, conviene que en este punto hagamos una precisión necesaria: una cosa es identificar conjuntos constructivos y otra muy distinta es decir que lo hacemos a través del “análisis clúster” de sus variables, como erróneamente deslizamos en la última de nuestras publicaciones sobre el tema (Azkarate 2010: 51, 55 y 61). En ninguna de las tres experiencias a las que nos hemos venimos refiriendo se hizo propiamente un análisis clúster, concebido como un conjunto de *técnicas multivariantes* con el objetivo principal de obtener mediante algoritmos, agrupaciones o *clusterings*.

Fue necesario que pusiéramos en marcha dos proyectos de mayor envergadura<sup>7</sup> para advertir, por primera vez, que la bondad de la idea original resultaba sin embargo insuficiente en contextos más complejos. Pronto fuimos conscientes de que tanto el análisis de la interrelación en el conjunto de variables y rasgos que suelen considerarse para caracterizar los objetos a analizar como el estudio de los aspectos sincrónicos y diacrónicos que están involucrados en estos procesos de variabilidad, constituían —tal y como había señalada G. Rolón en su tesis de

doctorado<sup>8</sup>— campos temáticos escasamente desarrollados (Rolón 2014: 2) y que, en consecuencia, resultaba necesario asumir nuevos retos metodológicos, aunque para ello tuviéramos que adentrarnos en ámbitos ajenos tradicionalmente a *l’archeologia del costruito*.

Es en este contexto de renovación metodológica en el que, desde GPAC y la Cátedra UNESCO *Cultural Landscapes and Heritage*, se intentará implementar procedimientos estadístico-matemáticos en los nuevos proyectos, incluyendo la automatización de algunos de los procesos. Constituyen buenos ejemplos de ello nuestras primeras propuestas sobre las potencialidades del láser escáner y las nubes de puntos (García-Gómez, Fernández de Gorostiza y Mesanza 2011) o las investigaciones llevadas a cabo por Noelia Osés (Osés y Azkarate 2013; Osés y Dornaika 2013; Osés, Dornaika y Moujahid 2014) sobre los sistemas expertos y la posibilidad de aplicar el análisis digital de imágenes a la extracción automática de contornos. Muy reseñable resulta también la tesis doctoral llevada a cabo por el arquitecto argentino G. Rolón a partir de una estancia de investigación en los laboratorios de GPAC, convirtiéndose en la primera investigación en la que se realiza un análisis estratigráfico y tipológico con apoyo de técnicas analíticas multivariantes para el estudio de edificios históricos (Rolón 2013 y 2014).

Pero será la tesis doctoral de uno de los firmantes de este trabajo la que ha supuesto, hasta la fecha, la apuesta más decidida en este sentido<sup>9</sup>. El objetivo de esta investigación fue la automatización de parte del proceso de lectura estratigráfica de alzados, empleando para ello el análisis multivariante, buscando con ello una mejora en la calidad, rapidez y productividad y eliminando parte de la subjetividad que puede acompañarle en muchos casos (Mesanza 2017: 79). Aunque resulte obvio, resulta necesario señalar que buena parte de los capítulos que siguen son deudores de este esfuerzo. Para este primer avance, se han reducido en lo posible las referencias explícitas a la metodología sobre algoritmos relacionados con las técnicas de minería de datos, que aparecerán en medios de publicación específicos.

<sup>7</sup> *La Cronotipología aplicada al estudio y conservación de los cascos históricos*, Ministerio de Educación y Ciencia. Dirección General de Investigación (HUM2007-65545/HIST). Plan Nacional de I+D+i (Azkarate *et al.* 2010) y *Montes de Vitoria: Análisis clúster aplicado al estudio estratigráfico de paisajes antropizados*. Proyecto de investigación financiado por GPAC (García-Gómez *et al.* 2011).

<sup>8</sup> *La vivienda popular riojana del ámbito rural: patrones arquitectónicos y contexto social en los valles durante el Periodo republicano*. Realizada por Guillermo Rolón y dirigida por Rodolfo Rotondaro y Adriana Callegari, fue defendida en la Universidad de Buenos Aires en abril de 2013.

<sup>9</sup> *Análisis multivariante aplicado a la lectura estratigráfica de alzados*, Tesis doctoral presentada por Amaia Mesanza Moraza, dirigida por Javier Gómez Lahoz (Universidad de Salamanca) y Agustín Azkarate Garai-Olaun (Universidad del País Vasco). Fue defendida en la Escuela Politécnica Superior de Ávila (Universidad de Salamanca) el 21 de julio de 2017, obteniendo por unanimidad la máxima calificación.

## LA IMPORTANCIA DE LAS TÉCNICAS CUANTITATIVAS

En consonancia con todo lo dicho, este artículo —que no es sino un avance de una obra más ambiciosa actualmente en curso— está concebido como un primer paso en el apasionante reto que la aplicación de las técnicas cuantitativas van a plantearnos en un futuro ya próximo. La automatización y las posibilidades que brinda la inteligencia artificial están ya entre nosotros y pronto tendrán protagonismo propio en la lectura estratigráfica de las construcciones históricas y, con carácter aún más general, en el conocimiento y gestión del Patrimonio Cultural y los Paisajes Culturales.

Habrán arqueólogos, arquitectos o historiadores del arte a quienes este tipo de propuestas no agrade particularmente por considerarlas demasiado frías y alejadas de los esquemas mentales que han acompañado tradicionalmente al estudio de las artes y las humanidades, pero los grandes avances que se han producido en los campos de la informática y de la documentación topográfica del patrimonio nos reafirman de algún modo en esta línea, abriéndonos una nueva puerta cuya sola mención asusta todavía a los más escépticos; nos referimos a la futura automatización de los análisis estratigráficos. ¿Podremos llegar algún día a una tecnología que asista efectivamente al arqueólogo en la no siempre fácil tarea de individualizar estratos en un muro? Estamos convencidos de ello, pero eso no sucederá si previamente no existe un procedimiento metodológico sobre el que se puedan fundamentar los primeros cimientos de un proceso que será sin duda complejo.

T. Mannoni se convirtió, una vez más, en fuente de inspiración. Fue relejendo su estudio sobre la evolución constructiva del castillo de Génova —un trabajo pionero para una época en la que aún ni la lectura estratigráfica de alzados, ni la propia AA habían nacido como tales— (Mannoni y Poleggi 1974), cuando descubrimos algunas claves conceptuales. Merece la pena citar en extenso sus propias palabras, pues evidencian con claridad que Mannoni era plenamente consciente de cómo un análisis matemático-estadístico básico podía revelar la existencia de paramentos producidos en momentos distintos:

Para conocer las técnicas constructivas [de la fábrica del “Castello”] tan solo hemos tomado en consideración los muros con despiece en hiladas, excluyendo todos los muros de yeso, es decir, aquellos realizados con técnicas más o menos “inciertas” y con materiales heterogéneos [...] hemos recurrido a dos factores distintivos, uno cuali-

tativo que define el tipo de elaboración de los sillares [...] y otro cuantitativo que refleja el despiece de los muros en función de la largura de los sillares, la altura de las hiladas y el espesor de las juntas. Para el segundo ha sido posible confeccionar un diagrama que tiende a evidenciar la existencia de eventuales “módulos”, conscientes o involuntarios, estos “módulos” pueden en parte depender de una cambiante disponibilidad de materiales, pero en cualquier caso pueden constituir elementos [indicios] cronológicos. Si consideramos las [dimensiones] medias de despiece de los sillares, apreciamos en los muros del Castillo dos agrupaciones bien diferenciadas y sin aparente continuidad entre sí. Al grupo con despiece de menor tamaño pertenecen todos los muros con técnica más tosca (juntas e hiladas irregulares y divididas [sdoppiate]), al otro, aquellos realizados con una técnica más depurada. Si [sobre el diagrama] a las medias de despiece (que han sido calculadas a partir de los sillares normales, es decir sin tener en cuenta aquellos con alturas y larguras anómalas), añadimos los valores de altura-largura máxima y mínima, apreciamos cómo los tres valores —“mínimo”, “medio” y “máximo”— representados coinciden casi siempre sobre una recta, lo que indica que existe una constante en la relación altura-largura, una relación que queda plasmada en la inclinación de la recta. La recta en sí puede considerarse por lo tanto una representación del módulo [del sillar] empleado en el muro. En lo que se refiere al grupo con despiece de menor tamaño, las mayores inclinaciones se aprecian en los muros realizados con técnica más tosca en los cuales han sido reutilizados materiales de la muralla prerromana y de desecho, mientras que disminuyen en aquellos muros que incorporan materiales simplemente despiezados y procedentes de frentes de cantera donde los estratos son delgados (1974: 186 y ss.).

Es muy interesante observar cómo el enunciado de Mannoni toca diversas cuestiones clave, cuestiones relativas no sólo a la importancia de registrar las dimensiones de las piezas que conforman los muros, sino también a la necesidad de tratarlas estadísticamente, ya que de ese modo pueden llegar a descubrirnos distintas realidades estratigráficas. Se nos ocurren, al hilo de las consideraciones de Mannoni, las siguientes observaciones:

- a) En el año 1974, no existía todavía un método estandarizado de “lectura estratigráfica de alzados”. Esto no quiere decir que Mannoni no apreciase la existencia de una estratificación en la estructura del castillo de Génova, sino que lo

hacia proponiendo una diferenciación según técnicas constructivas, es decir, según un análisis tipológico sistemático. En definitiva, Mannoni no hacía otra cosa que recurrir al acervo corriente en la arqueología de entonces: es consciente de la existencia de estratos y sin embargo su forma de individualizarlos no se fundamenta en la “estrategia interfacial” que se impuso *a posteriori*, sino en la “estrategia tipológica” que era la tradicionalmente empleada en nuestra disciplina prácticamente desde sus orígenes.

- b) Un detenido repaso del artículo en su conjunto revela que el sistema de registro empleado por Mannoni se encuentra altamente sistematizado; tanto la descripción de los muros que él denomina “cualitativa” como la que califica de “cuantitativa” encajan en lo que podríamos definir como una descripción tipológica basada en la consideración por separado de un cierto número de variables. Variables que tienen en cuenta —entre otras— características como la granulometría de la arena de los morteros, la presencia porcentual de ciertos materiales como el cuarzo, las ofiolitas o la caliza marmórea también en los morteros, la tipología de los materiales reutilizados, el tipo de talla, etc., aparte por supuesto de las dimensiones de los sillares, de los que se mide la altura, la largura, y el grosor de las juntas.
- c) Como puede apreciarse, Mannoni ensaya el “análisis clúster” de algunas de las variables registradas. Un análisis que destaca por el incipiente empleo de un procedimiento matemático-estadístico, en el cual se observan: a) la preselección de una población de estudio (sólo los muros con despiece en hiladas); b) la desestimación de valores anómalos u *outliers*<sup>10</sup>; c) la identificación de patrones o “módulos”; d) el cálculo de “medias”; y e) su reflejo en distintos diagramas de dispersión.

## PUESTA EN PRÁCTICA DE UN EJEMPLO CONCRETO

Como sin duda el lector estará comprobando, la problemática que estamos intentando abordar encierra una

gran complejidad, y no sólo en un sentido práctico, sino también en un sentido teórico. Es por ello que, antes de sumergirnos en una descripción más pormenorizada del procedimiento que hemos seguido, vamos a presentar un pequeño esquema descriptivo del mismo, no sin antes hacer un par de apuntes.

Uno primero para recalcar que, si bien el eje vertebrador del procedimiento que hemos diseñado sigue siendo —como en nuestros anteriores trabajos— el análisis clúster, su protagonismo se ve ahora bastante matizado por un nutrido conjunto de procesos colaterales que son claves si en un futuro queremos lograr una automatización efectiva del proceso de lectura estratigráfica.

El segundo apunte es en realidad un “esquema del esquema” porque creemos que conviene advertir que el procedimiento puede resumirse aún más en tres grandes fases de trabajo; una primera, que tiene que ver con la captura de los datos geométricos, es decir, con la medición de los elementos que componen el edificio objeto de estudio (o si se prefiere, con la cuantificación de la realidad material); otra segunda, que se refiere propiamente al tratamiento matemático-estadístico de los datos obtenidos en aquella fase previa; y una última, que tiene que ver con la plasmación de los resultados obtenidos sobre un modelo del edificio analizado y el hallazgo de la secuencia estratigráfica que caracterizó su evolución constructiva.

Pues bien, ahora ya más en detalle, creemos que nuestro procedimiento puede sintetizarse en ocho pasos, que serían:

1. Elección de variables.  
Como hacia el final del epígrafe sobre captura y extracción de datos geométricos explicamos, para este trabajo hemos decidido trabajar exclusivamente con variables geométricas, es decir, midiendo aquellas distancias que son descriptivas de la forma que tienen los objetos. Hemos decidido que la unidad base de nuestro análisis van a ser los sillares que componen el edificio, por lo que los hemos medido uno a uno.
2. Captura geométrica.  
Una vez teníamos claro qué es lo que teníamos que medir, el siguiente paso fue emplear el instrumental topográfico adecuado, en este caso el láser-escáner, para lograr una medición lo más precisa posible.
3. Delineación.  
El producto primario que nos proporcionaba la medición por medios topográficos era un modelo de nube de puntos que no podía ni puede emplearse

<sup>10</sup> Según el texto citado: “las medias de despiece, que han sido calculadas a partir de los sillares normales, es decir sin tener en cuenta aquellos con alturas y larguras anómalas”.

directamente para el tratamiento estadístico, por ello fue necesario un tratamiento secundario o “delineación” que nos sirvió para la definir el contorno de los sillares mediante polilínea.

4. Extracción de datos geométricos.  
Una vez delineados todos los sillares que componen el inmueble, procedimos a la programación de un conjunto de rutinas que extrajesen de ese modelo lineal todos los datos que precisábamos en función de las variables definidas en el punto 1.
5. Análisis estadístico predictivo.  
Para poder aplicar el análisis clúster a un conjunto de elementos es preciso que todos estos tengan características similares. Las dovelas de los arcos o las piezas que presentan algún recorte a modo de engatillado, aun siendo sillares, no pueden ser objeto de análisis clúster conjunto. Para filtrarlas y sacarlas fuera del grupo de trabajo es para lo que empleamos las técnicas estadísticas de carácter predictivo.
6. Análisis clúster.  
Una vez depurada la muestra, fueron los sillares restantes —que seguían siendo la inmensa mayoría de los que componen el edificio— los que sometimos al análisis clúster. Este análisis, en un primer momento, consiste en la determinación del número óptimo de clústeres en que puede dividirse dicha muestra y en un segundo momento en el análisis clúster propiamente dicho, que es el que nos devuelve las agrupaciones de sillares que forman parte de una misma unidad.
7. Mapeado de clústeres.  
Del mismo modo que los pasos 1 a 4 nos sirvieron para pasar de la realidad material de las cosas a esa síntesis esencial que son los datos cuantitativos que subyacen en ella, una vez obtuvimos los clústeres, tuvimos que hacer el “viaje de vuelta” y trasladarlos al espacio físico real para comprobar con su representación sobre el edificio si podía subyacer una lógica estratigráfico-constructiva en su distribución.
8. Deducción de la secuencia estratigráfica.  
Al estudiar la distribución espacial de los clústeres sobre los edificios, detectamos que los sillares que conformaban un mismo clúster presentaban sobre los muros una tendencia a la concentración que sugería efectivamente que nos encontrábamos ante una realidad estratigráfica, una realidad que podría estar informándonos sobre la secuencia constructiva del edificio, si tratábamos de leerla según los principios de estratificación arqueológica.

Insistamos a propósito de este esquema que éste no es sino una síntesis provisional que, de hecho, como resultado del proceso de experimentación que a continuación vamos a exponer, ya estamos viendo que va a tener que ser alterado en alguno de sus puntos (ver apartado conclusiones).

## Elección del objeto de estudio

Volvamos de nuevo al inicio y comentemos cómo nuestro trabajo comenzó en el momento mismo de la selección del edificio que queríamos estudiar arqueológicamente. La edificación elegida fue la Ermita de la Purísima Concepción, construida a finales del siglo XII, y ubicada en la localidad treviñesa de San Vicentejo, a 14 km al sur de Vitoria-Gasteiz. Para muchos investigadores se trata de un templo “excepcional” dentro del panorama del Románico tardío peninsular debido fundamentalmente a sus extraordinarios valores formales (Castiñeiras 2012: 231).

Esta elección, a primera vista, podría parecer sorprendente dada su aparente sencillez y homogeneidad además de su reducido tamaño; pareciera que hubiéramos querido elegir una construcción “fácil” donde contrastar nuestras hipótesis, aunque nada más lejos de la realidad. En esta decisión se tuvieron en cuenta aspectos muy diversos que pasamos a detallar brevemente.

Un primer punto que se consideró es la inexistencia de estudios arqueológicos en profundidad realizados sobre la ermita. Investigadores de la talla de Francisco Iñiguez Almech (1968: 203-207) o Micaela Portilla (1968: 188-193) han estudiado el templo, pero desde el punto de vista de la Historia del Arte y no tanto desde una perspectiva de la Arqueología de la Arquitectura. Este hecho permite que nuestra posición de partida sea neutra, sin ninguna influencia que incline la investigación hacia uno u otro lado.

El segundo aspecto a destacar es la aparente homogeneidad de los paramentos de la iglesia, circunstancia ésta que, en contra de lo que pudiera parecer, suponía un reto aún mayor al poder plantear, como hipótesis de trabajo, que la metodología de lectura estratigráfica automatizada permitiría distinguir anomalías que con el modo tradicional de proceder seguramente hubieran pasado desapercibidas.

En último lugar destacaremos la litología de sus paramentos, ya que en todos ellos aparece como roca preferente para su construcción las calizas o lumaqueles de Ajarte, trabajadas mayoritariamente en forma de



Figura 1. Vista panorámica de la ermita de la Purísima Concepción, San Vicentejo (Condado de Treviño, Burgos).

sillar, un aspecto bastante habitual en otros templos de la geografía alavesa. Esta cuestión es importante para la reproducibilidad del experimento. Esto no quiere decir que el método que aquí planteamos solo funcione con fábricas realizadas con lumaquela de Ajarte, ya que el método en sí mismo es extrapolable a cualquier inmueble, pero no los valores absolutos que en él se obtengan.

Tal y como señalaba Mannoni en su proyecto, la muestra con la que se había de realizar el análisis clúster debía ser homogénea, eliminando aquellos valores anómalos presentes en la misma. Esto requiere de una serie de trabajos previos para “depurar” la base de datos que en nuestra investigación solventamos aplicando diferentes técnicas predictivas (como son el “análisis discriminante” o los “árboles de clasificación”). Eso supuso tener que recopilar datos de un conjunto de edificios que servirían para el aprendizaje autónomo del sistema, y otros diferentes donde poder aplicar dicho conocimiento. En el currículo de nuestro grupo de investigación (GPAC) disponíamos de una amplia colección de ellos, repartidos por toda la Comunidad Autónoma del País Vasco, pero optamos por explorar el entorno alavés en busca de aquellos que más similitudes tuvieran con la ermita románica previamente elegida. Finalmente, nos

decidimos por la Basílica de Armentia y por la Catedral de Santa María, dos de los templos más emblemáticos del territorio de Álava y objeto ambos de sendos estudios de investigación por nuestro grupo.

### Captura y extracción de datos geométricos

Aunque previamente ya se tenían ciertas previsiones sobre las técnicas de documentación geométrica que podríamos utilizar en la investigación, no queríamos, sin embargo, descartar ninguna posibilidad antes de comprobar en la práctica cuáles eran las ventajas/desventajas de una u otra metodología de registro. Inicialmente ensayamos con los métodos fotogramétricos (con fotogrametría convergente principalmente), pero pronto observamos que para llegar a las precisiones que íbamos a requerir, necesitábamos tal número de capturas fotográficas que, el tiempo para obtenerlas (y la inevitable acumulación de errores que conlleva encajarlas en un único modelo), iba a constituir un verdadero obstáculo para la viabilidad del proyecto. Las pruebas, en cambio, que hicimos con el láser-escáner fueron incomparablemente más satisfactorias, razón por la que finalmente nos decidimos por este último sistema.

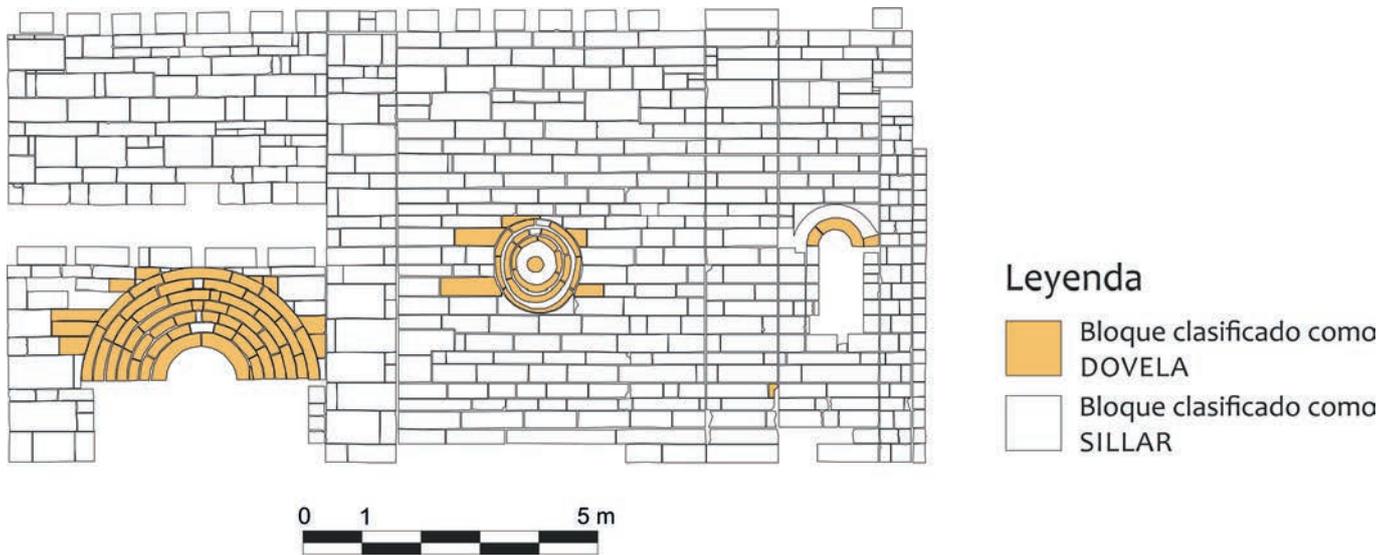


Figura 2. Alzado exterior sur de la ermita de San Vicentejo con las dovelas de arco identificadas automáticamente. Aunque no era objeto de este artículo, vemos cómo es factible la búsqueda de determinados elementos aplicando diversas técnicas predictivas.

Para que el lector no especialista pueda hacerse una idea de la exhaustividad de las mediciones efectuadas, diremos que el modelo de nube de puntos obtenido una vez completados los escaneos en San Vicentejo, se compone de más de 12 millones de puntos, medidos con una resolución que oscila entre los 5 y los 10 mm. Estos modelos están compuestos por constelaciones de puntos que se encuentran aislados unos de otros en el espacio, siendo necesario unirlos para representar realidades geométricamente significativas.

En la medida en que —como pronto explicaremos— la realidad que nos interesaba capturar era la de la dimensión de los sillares, tuvimos que someter esos millones de datos espaciales a un post-proceso que permitiera seleccionar y unir en cada caso aquellos puntos que conformaban el perímetro de dichos elementos constructivos; este post-proceso exigió que un operador humano —asistido por el pertinente software CAD— delinease mediante una polilínea el contorno de cada bloque pétreo. Se trata de una tarea ardua más que complicada, una tarea que hoy por hoy —aunque distintos fabricantes ya están trabajando en el desarrollo de un algoritmo para la automatización del delineado— no se puede realizar sino manualmente. Del delineado efectuado han resultado algo más de 6800 polígonos que conforman el modelo definitivo sobre el que se ha efectuado el análisis matemático-estadístico, una vez que se introdujeron los datos referidos a cada bloque en un Sistema de Información Geográfica. Se programaron para ello sencillas rutinas que permitieran una extracción automática de la información referida y su

almacenamiento en una base de datos creada a su vez en el ya citado SIG. Con todo, en la medida en que las herramientas de análisis que incorporan los SIG no están específicamente pensadas para buscar —mediante procedimientos estadísticos— aquellos patrones ocultos en los datos, tuvimos que exportar nuestra base de datos a programas propios para la llamada “minería de datos”.

Antes de continuar conviene que nos detengamos, sin embargo, en un punto importante al que nos hemos referido brevemente en el párrafo anterior. Todo método de reducción de la realidad a números —y eso es lo que necesitamos para poder hacer cálculos estadísticos con ellos— pasa siempre por algún tipo de sistema de medición; se mide el peso, se mide la intensidad de la luz o del sonido, se mide también la velocidad, la acidez de un fluido, la población de una ciudad o las tendencias de opinión de un país. En el caso de nuestra experiencia de San Vicentejo, lo que cuantificamos es la espacialidad de las cosas; medimos posiciones y distancias con las que podemos describir la forma y disposición de las piezas que componen el edificio.

Al decantarnos por la medición de distancias no estamos diciendo que para realizar un análisis clúster no sea útil la medición de otras características materiales del edificio como litología, patología, marcas de cantero, talla, labra, etc. Este fue, de hecho, nuestro proceder metodológico en las diversas experiencias de investigaciones que hemos mencionado en los capítulos precedentes y en las que comenzábamos identificando —como en el caso de la Catedral de Santa María— *variables tanto de carácter técnico-constructivo*

(tipos de materiales constructivos, tipos de aparejos, tipos de talla, marcas de cantero, etc.) como *variables de carácter formal* (tipos de puntillas de los arcos trilobulados y del antepecho del triforio, tipos de capiteles, etc.).

Pero la cosa cambia radicalmente cuando se adentra uno en el complejo mundo de la estadística, los análisis multivariantes y la automatización. Es por ello por lo que, en este nuevo reto y en la medida en que estamos emprendiendo una labor pionera en busca de una mayor sistematización de los métodos de análisis estratigráfico aplicados en Arquitectura, hemos optado porque el trabajo sea lo menos “intuitivo” (o lo más “cuantitativo” si se prefiere) posible, razón por la cual hemos huido de las variables “cualitativas”. Esto es lo que explica que, por ahora, en esta propuesta de método, nosotros prefiramos trabajar exclusivamente con datos de la medición de distancias y posiciones en el espacio.

### Clasificación (Análisis clúster)

Así pues, descartados todos los elementos diferentes a un sillar perfecto, ya disponíamos de una serie susceptible de someterse a un análisis clúster, ya que sus componentes eran lo suficientemente parecidos y al mismo tiempo lo suficientemente distintos, como para establecer entre ellos una comparativa coherente y significativa desde el punto de vista tipológico.

#### *Determinación de las variables estadísticamente más significativas*

Probablemente en un futuro (cuando contemos con una base de datos de edificios analizados más amplia) podamos saber, ya de antemano, qué variables van a ser estadísticamente significativas para nuestro análisis; de ese modo podremos ahorrar mucho tiempo a la hora de tener que tomar mediciones en un inmueble. Por ahora sin embargo nos vemos obligados a seguir siendo maximalistas, es decir, a ser exhaustivos, y considerar, por ejemplo, al medir la cara de un sillar, su alto, su ancho, sus dos diagonales, la posición de su centro, la longitud de su perímetro, su área, la de su envolvente, etc. Por el momento, sólo después de considerar tantas mediciones como nos sea posible, será la propia estadística la que analizando la naturaleza matemática de esos datos pueda decirnos cuáles de esas mediciones son verdaderamente significativas al objeto de diferenciar estratos.

En el caso de las medidas relativas a los sillares perfectos de San Vicentejo, sometimos los datos a un análisis de componentes principales (al que coloquialmente los expertos suelen referirse sencillamente como PCA). No es este el lugar para entrar en los principios matemáticos sobre los que se fundamenta este procedimiento; no obstante, puede decirse que consiste en transformar el conjunto inicial de variables en otro mucho más reducido, pero de forma que reflejen un porcentaje bastante amplio de la variabilidad total de la muestra (Pérez López y Santín González 2007: 351).

Como en todo procedimiento indiciario, la distinción de estratos parte de un principio básico que es la detección de la anomalía, o, dicho de otro modo, la detección de homogeneidades y heterogeneidades. Así, dos o más sillares iguales presentes en un mismo contexto deberían en un principio entenderse como pertenecientes a una misma realidad, mientras que dos sillares distintos deberían normalmente adscribirse a dos realidades diferenciadas.

La elección de las variables es un aspecto clave en todo el proceso, ya que se han de elegir aquellas que dispongan de información suficiente como para poder discriminar los distintos grupos que se formen. Puede ocurrir que al trabajar con una sola variable no se aprecie ningún tipo de estructura interna pero que la combinación del mismo atributo con algún otro produzca una salida completamente diferente. En cualquier caso, no se trata de disponer de infinidad de variables, ya que un exceso de las mismas puede llevar a un enmascaramiento del fenómeno, sino de buscar aquellas que más puedan aportar al proceso.

Para nuestra investigación, de las 33 variables con las que inicialmente se definió cada bloque de piedra, finalmente, para este análisis de conglomerados, optamos por trabajar únicamente con el alto y largo de piedra, al igual que lo haría Mannoni en su proyecto del castillo de Génova y que tan buenos resultados le reportó.

#### *Determinación del número de clústeres*

Una vez habíamos llegado a la conclusión estadística de que las variables “largo” y “alto” eran las más significativas, ahora debíamos servirnos de ellas para concretar cuáles eran los clústeres presentes en la fábrica de San Vicentejo. Pues bien, los clústeres se determinan por así decir en dos tiempos; primero, se procesan los datos sólo con objeto de determinar el número de clústeres en que éstos parecen compartimentarse, y ya después, conociendo

esa cifra, se hace un nuevo procesado para determinar concretamente qué elementos —en nuestro caso, qué sillares— pertenecen a un clúster y cuáles a otro.

Vayamos en orden; para conocer el número de clústeres que conforman nuestra muestra, hemos empleado dos procedimientos. En efecto, aunque en principio no es necesario emplear más de uno (existen varios posibles), queríamos asegurarnos de que, independientemente del procedimiento estadístico que escogiésemos, los resultados obtenidos iban a ser los mismos. Así por un lado, hicimos una primera comprobación mediante el llamado “método del codo” (“ELBOW method” en su expresión anglosajona) que lleva ese nombre precisamente porque la gráfica resultante de los cálculos que implica consiste justo en eso, es decir, en una línea que llegado cierto punto gira bruscamente; pues bien, ese punto de inflexión corresponde con un cierto valor que es el que indica el número óptimo de clústeres que explican nuestra muestra, el cual en este caso, ascendía a un total de cinco clústeres.

El segundo método empleado para determinar el número de grupos en que se dividía nuestra muestra, consistió en la confección de dos histogramas, uno con datos de la variable “alto” y otro con datos de la variable

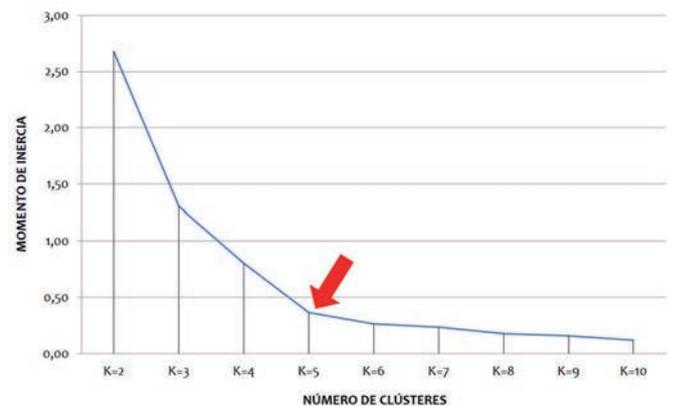


Figura 3. Gráfico resultante de aplicar el contraste estadístico Elbow en los alzados de la ermita de San Vicentejo. Se aprecia como a partir de  $k=5$ , la gráfica hace un quiebro, punto que señala el número de agrupaciones más probable de la muestra.

“largo”. Comparativamente, mientras que en la gráfica de las larguras no se distinguían nítidamente repuntes o modas que indicasen que las medidas tendían a agruparse en torno a ciertos valores, en el caso del histograma de la altura esas modas aparecían bien destacadas con sus dientes de sierra concentrados fundamentalmente en cuatro puntos, al que —siguiendo a Yan (2005)— cabría añadir un quinto, conformado por los llamados valores atípicos u “outliers”.

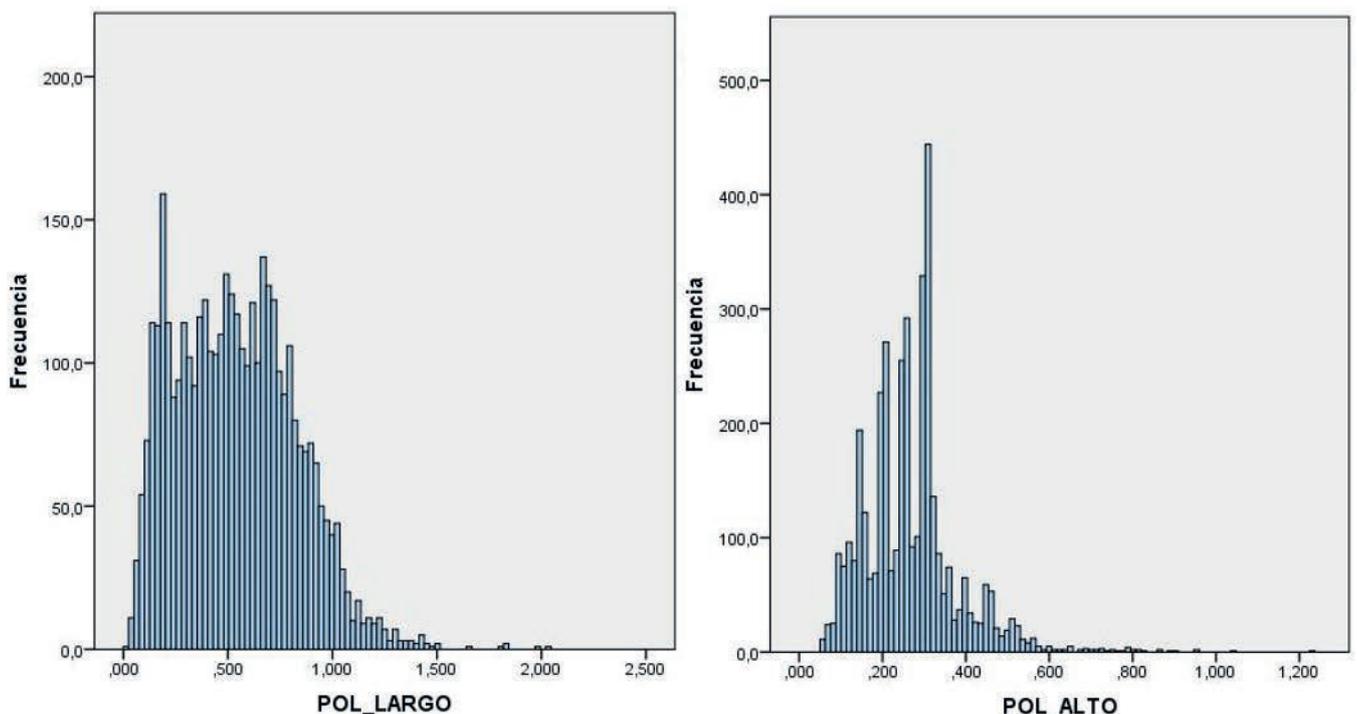


Figura 4. Histogramas realizados a partir del largo y alto de los sillares. Mientras que en el primero de ellos apenas se reflejan picos significativos, en el segundo se aprecia la presencia de al menos 4 modas, lo que puede estar indicando la presencia de 4 grupos diferentes. (Histogramas realizados con el paquete estadístico SPSS a partir de 3.885 elementos. Para la variable largo:  $\text{media}=0,540 \text{ m}$ ;  $\sigma = \pm 0,285 \text{ m}$ , mientras que para el alto:  $\text{media}=0,269 \text{ m}$ ;  $\sigma = \pm 0,113 \text{ m}$ ).

Tomando por lo tanto esos cinco como número óptimo de clústeres en que puede dividirse nuestra muestra, ya podíamos dar el siguiente paso para definir concretamente qué sillares podrían agruparse dentro de un conglomerado u otro. Existen para ello diversos algoritmos que, según trabajen con la “moda” o con la “media” de los datos se denominan “*k-means*” o “*k-modal*”. Por la naturaleza de nuestros datos, en principio cualquiera de los dos podía sernos útil, si bien finalmente optamos por “*k-means*” dado que es el más extendido entre los distintos softwares de procesamiento estadístico.

### Mapeado de los clústeres

Una vez aplicado el “*k-means*” el sistema conformó los clústeres indicándonos sillar por sillar, cual pertenecía a una agrupación y cual a otra. Ahora bien, a pesar de ello, el hecho de que esas agrupaciones fueran coherentes matemáticamente según los datos suministrados, no implicaba necesariamente que esa coherencia fuera también arquitectónica o estratigráfica, de modo que lo siguiente que tuvimos que hacer para comprobarlo fue mapear los clústeres obtenidos mediante el SIG y ver cuál era su distribución tridimensional en la iglesia de San Vicentejo.

Una vez plasmados sobre el modelo, los clústeres definidos estadísticamente presentaban una distribución a la que no estamos habituados los arqueólogos cuando hacemos una lectura estratigráfica convencional. Sin embargo, a nuestro juicio este hecho, lejos de constituir una dificultad, supone un verdadero reto científico ya que introduce interesantísimas cuestiones relativas no sólo a la verdadera naturaleza de la estratificación sino sobre todo a nuestra capacidad real para percibirla. Los estratos que identificamos sobre los muros mediante una lectura de alzados convencional tienden a estar bien localizados, a ser compactos, a tener un contorno bien definido; la pregunta sin embargo es ¿son verdaderamente así, o su forma depende —al menos en parte— del modo en que nos han enseñado a visualizarlos? Se trata sin duda de una cuestión apasionante a la que estamos lejos aún de poder dar una respuesta definitiva. Con todo, en el caso de San Vicentejo hemos optado —obviamente— por aquella vía que creemos más certera.

Atendiendo al modo en que se distribuyen por el edificio, nuestros cinco clústeres pueden clasificarse en dos apartados; por un lado, estarían aquellos completamente dispersos a los que resulta muy difícil atribuir una entidad estratigráfica (aunque en realidad este sólo



Figura 5. Desarrollo de los alzados interiores y exteriores de la ermita de San Vicentejo y plasmación sobre ellos de los clústeres calculados.

es uno, el clúster 2), y por otro aquellos otros que se presentan claramente concentrados en ciertas partes del edificio y que sí parecen responder a un hecho estratigráfico. Simplificando al extremo, el estudio visual del modelo 3D de San Vicentejo con los sillares coloreados según el clúster al que pertenecen, apunta a que:

1. El clúster 1 se concentra en las bóvedas, tanto en la de las naves como en la del testero.
2. El clúster 2 no se aprecia nucleado en ninguna zona en concreto.
3. El clúster 3 se aglomera en los muros del primer tramo de nave, tanto al interior como al exterior.
4. El clúster 4, aunque bastante disperso, se encuentra notablemente concentrado en la parte superior del paramento exterior del citado cierre.
5. El clúster 5 se encuentra zonificado en los contrafuertes, en las hiladas superiores de los muros laterales, en la mayor parte del paramento interior del cierre occidental, y en la parte inferior del paramento exterior de ese mismo cierre.

No obstante, como es evidente, aunque aceptemos que los clústeres 1, 3, 4 y 5 son estratos, esto por sí sólo no es suficiente para poder conocer la secuencia constructiva de un edificio. Antes hay que estudiar las relaciones físicas que median entre ellos para poder deducir una secuencia temporal de anteroposterioridad.

## **FUTURO DESARROLLO. HACIA UNA CONSTRUCCIÓN-CLÚSTER DE LAS SECUENCIAS ESTRATIGRÁFICAS ARQUITECTÓNICAS**

Parece evidente que las concentraciones de sillares de un mismo clúster —en tanto que “distribuciones no aleatorias”— deberían ser síntoma de “localizaciones discretas”, es decir, de localizaciones concretamente escogidas para el desarrollo de una acción, mientras que las “distribuciones aleatorias” no serían producto de acción concreta alguna sino del puro azar. En consecuencia, parece lógico deducir que aquellas partes del muro en que se aprecia la concentración de un mismo clúster, son resultado de ciertas acciones constructivas concentradas, no sólo en un mismo espacio, sino también en un mismo lapso de tiempo. Dicho de otro modo, creemos que esas concentraciones son, si no estratos directamente, al menos sí, sus núcleos fundamentales.

El principio sobre el que hemos basado el anterior razonamiento no es evidentemente nuestro, lo hemos tomado del clásico de Hooder y Orton, *Análisis espacial en arqueología* (1990: 41). Los análisis de distribuciones han sido cada vez más frecuentes en arqueología, siendo su principal objetivo precisamente el de identificar patrones no aleatorios que pudieran delatar la presencia de la acción humana partiendo de la idea de que una distribución espacial notablemente aleatoria apunta a una ausencia de dicha acción humana, mientras que la existencia de aglomeraciones sugiere la presencia de un patrón de comportamiento racional.

## **De la identificación de estratos hacia la secuencia constructiva**

Nótese al respecto cómo, del mismo modo que el análisis de la estratificación de un edificio no acaba con la identificación de estratos, tampoco la identificación de clústeres con valor estratigráfico puede servirnos por sí sola para completar la lectura de alzados de un edificio. Es preciso ir un paso más allá hasta lograr reconstruir una secuencia cronológica coherente y para ello, como es obvio, es preciso analizar el tipo de relaciones físicas que median entre los estratos identificados: sólo a partir de ellas se puede deducir la coordenada temporal y a su vez de ésta la secuencia constructiva del conjunto.

Curiosamente, los arqueólogos estamos tan acostumbrados a trabajar con esas relaciones estratigráficas que no siempre somos conscientes de los signos externos en que nos hemos fijado para determinar qué hechos estratigráficos se han producido antes y cuáles se han producido después. Sin embargo, tomar consciencia de cuáles son estos signos resulta clave si, como es nuestro caso, lo que pretendemos en un futuro es poder automatizar su detección.

Empecemos señalando cómo todas las relaciones estratigráficas vienen de un modo u otro denotadas por uno o varios de estos signos, y a su vez éstos son claramente identificables por su distinta morfología y/o posición en el espacio. En consecuencia, si se cumple el requisito de haber hecho una buena documentación topográfica del edificio, no cabe duda de que alguna forma algorítmica habrá para detectar dichos signos partiendo de los datos recogidos en nuestro repositorio. El problema radica evidentemente en dar con esas formas algorítmicas.

Nuestro experimento en San Vicentejo se ha ceñido fundamentalmente al aparejo de sillería, una técnica



Figura 6. Esbozo a grandes trazos de la evolución histórico-constructiva de San Vicentejo, realizada a partir de la localización sobre los alzados de los 4 clústeres más significativos.

constructiva en que los signos externos que pueden ser indicativos de antero-posterioridad son bastante concretos:

1. El indicio más básico lo observamos cuando un estrato se sitúa todo él encima de otro; el estrato de arriba será en principio más moderno que el de abajo. Pues bien, la forma de automatizar la detección de esta relación estratigráfica consiste en este caso en enseñar al sistema a discernir —de la geometría absoluta del modelo de San Vicentejo— qué clúster es el que está en la cota más alta y cuál en la más baja, para de ese modo colegir sus respectivas cronologías relativas.
2. Otros indicios relativamente básicos son aquellos que denotan que para colocar un estrato ha habido que alterar al menos una parte de la superficie de otro. Normalmente el estrato que ha visto alterado su aspecto original será el más antiguo y el que no muestra esos signos será el más moderno. En este caso, para que el sistema aprenda a identificar tales tipos de relación, éste deberá ser capaz de detectar los signos de tal remodelación; tratándose de sillería, el más evidente lo conforman los llamados “engatillados” y éstos pueden identificarse algorítmicamente.
3. Más compleja es la cuestión de aquellos estratos que se sitúan uno junto a otro sin que se observen abruptas alteraciones en ninguno de ellos. Se trata de los casos en los que los estratos están unidos/separados por una interfaz vertical limpiamente recta. En tales circunstancias, un primer paso, consiste evidentemente en detectar esa interfaz vertical, un paso que es bastante factible de automatizar, si bien el segundo ya no lo es tanto, porque sería necesario enseñar al sistema a que recabase nuevos datos indirectos como el de que esa interfaz vertical corresponde a la línea de jamba de una ventana que ha sido rellenada o que se trata sin embargo de la antigua esquina de un edificio que luego ha sido ampliado, entre otros casos posibles.

En cualquier caso, a partir de estos tres tipos de indicios se pueden deducir prácticamente todas las relaciones estratigráficas, y lo más interesante de todo es que la automatización de su detección resulta *a priori* bastante factible. De hecho, la individualización de engatillados ya la ensayamos al principio del procesamiento de los datos de San Vicentejo (aunque entonces sólo la hicimos con el fin de descartar los anómalos sillares que los contienen). Por su

parte la detección de interfaces verticales también hemos logrado automatizarla. En ambos casos hemos programado el SIG para que resalte su posición dentro del modelo, con objeto de que —en tanto en cuanto no logremos una verdadera automatización en la detección de las relaciones de anteroposterioridad— al menos el operador/arqueólogo pueda valerse de ellas para determinar la secuencia clúster/estratigráfica del edificio que está analizando.

### **Secuencia constructiva detectada automáticamente en San Vicentejo**

Aunque como ya hemos apuntado, todavía nos encontramos en proceso de desarrollar un método automatizado de detección de relaciones estratigráficas, y por lo tanto éste aún no ha podido ser aplicado en San Vicentejo, otro indicio que apunta a que efectivamente los clústeres detectados en esta pequeña iglesia dan la sensación de ser realidades estratigráficas tal como habitualmente las concebimos en arqueología, es el del modo en que los clústeres se encuentran dispuestos —unos con respecto a otros— dentro del edificio; este modo no contradice, sino que se presta bastante bien a una lectura coherente del inmueble desde el punto de vista del proceso constructivo. Ahora bien, es preciso advertir, que cuando trasladamos los datos procesados algorítmicamente al modelo tridimensional del edificio estudiado (Fig. 5), éstos no muestran ese aspecto nítido y perfectamente contorneado al que estamos acostumbrados cuando vemos la representación gráfica de una lectura de alzados. Es por esta razón que —mientras vamos depurando nuestros métodos en busca de unos resultados más parecidos al de los estándares comúnmente aceptados, y sólo con el fin de que se comprenda más fácilmente la lógica estratigráfica que, estamos convencidos, subyace en nuestros datos— hemos elaborado una figura (Fig. 6) que, partiendo de los mismos datos en bruto, representa la estratigrafía tendencialmente, es decir, atendiendo principalmente a la concentración de los clústeres y no a una definición perfecta de sus contornos.

Pues bien, hecha ya esta advertencia, empezaremos diciendo que los datos derivados del análisis clúster sugieren que el edificio empezó a construirse por el ábside y que en esa primera fase de obras se completaron también los muros laterales correspondientes al primer tramo de nave, incluyendo los contrafuertes exteriores que median entre éste y la cabecera. Además, en el caso del paño septentrional, los sillares que componen este

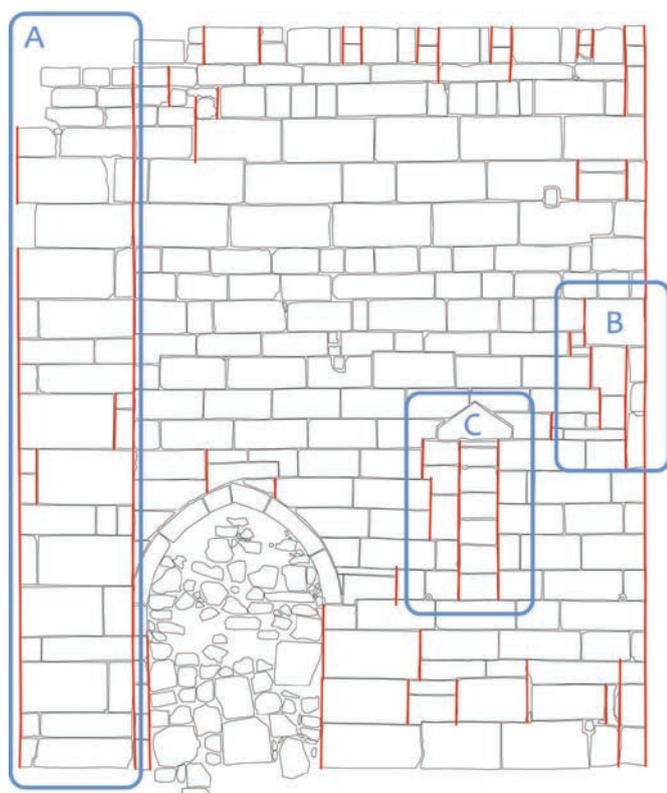


Figura 7. Identificación sobre los alzados de las interfaces verticales, programada para obtenerse de una manera automatizada, sin intervención del investigador. En la figura aparecen diferentes situaciones que podemos encontrar: la zona A muestra un ejemplo de la delineación de las aristas de los contrafuertes; en la B es un ejemplo de interfaz constructiva; y en la C las líneas se corresponden en parte con el cegado del vano y en parte con el corte realizado para encajar la ventana.

clúster alcanzan la parte inferior del muro del segundo tramo de nave lo que indica que una parte de éste también se construyó en esta fase. Lo que se observa al exterior para este clúster, se repite más o menos igual hacia el interior, salvo que aquí se observa —mejor que por fuera— el hecho de que la parte inferior del muro meridional del segundo tramo de nave también pertenece a esta fase. Huelga añadir para completar esta descripción, que en este momento se construyen también los cuatro pilares que separan los distintos tramos de nave.

En la segunda fase, apoyada sobre la obra anterior, se erige la bóveda de cañón, hecha toda ella a base de sillares de reducidas dimensiones.

En una tercera fase, los artífices parecen querer cerrar por completo la volumetría del templo que, todo apunta, había quedado inacabada en algunas zonas. Nos estamos refiriendo fundamentalmente al muro del hastial occidental y la parte superior del segundo de los muros del segundo tramo de naves. En el primer tramo de naves se produce asimismo un recrecido de varias

hiladas, probablemente para buscar mejor acomodo a un tejado que ahora debería tener en cuenta y respetar la altura de la bóveda.

Finalmente, la cuarta fase se explica, además de por la distribución del clúster en sí, por el hecho de que la morfología de los sillares que lo componen (se trata más de losetas que de sillares propiamente dichos) indica que hablamos de una última reforma o acabado al que se somete a algunas de las hojas exteriores de los muros. Estas modificaciones, más bien superficiales, se concentran en varias zonas: en todo el paramento interior del muro de cierre del hastial occidental; en la parte inferior del paramento exterior de ese mismo muro; en los contrafuertes exteriores que separan el primero del segundo tramo de naves (lo que nos hace sospechar que el núcleo de estos machones puede en realidad haber sido fabricado en alguna fase anterior); y también en los paramentos exteriores de los muros septentrional y meridional, en una franja de hiladas situadas en la zona superior de los paños, un poco por debajo de la altura de los canecillos.

## CONCLUSIONES

El futuro es bastante prometedor porque aún estamos muy lejos de agotar todas las vías estadístico-matemáticas y geométricas con fines de detección estratigráfica. En el caso de San Vicentejo no hemos alcanzado las cotas de automatización a las que creemos que se puede aspirar siendo realistas, pero cuando menos contamos con un prototipo de asistente a la lectura de alzados que brinda al arqueólogo la posibilidad de llegar a visualizar esa estratigrafía subyacente tan difícil de detectar por medios convencionales.

Por nuestra experiencia en esta investigación, tenemos claro que el campo que se abre —aunando Arqueología de la Arquitectura, documentación geométrica por topografía y técnicas de minería de datos y análisis estadístico— está lleno de posibilidades. Qué duda cabe que el nuestro es sólo un pequeño primer paso, pero éste ya nos permite intuir que el camino que se abre apunta hacia la inteligencia artificial y los sistemas expertos. A nadie le resulta extraño hablar de inteligencia artificial en el campo de la ingeniería, de las telecomunicaciones, de la visualización 3D e incluso de la medicina.

Los sistemas expertos —SSEE— son una rama de inteligencia artificial —IA— y fueron desarrollados a mediados de los 60. Un sistema experto no es nada más que un sistema informático al que se le provee de

información en un campo específico de conocimiento y al cual se programa para imitar los procedimientos y decisiones de los expertos (Barceló 2010: 13). Proporcionan soluciones a una gran variedad de problemas que de no ser por ellos no podrían ser resueltos empleando los métodos más tradicionales. Dentro de estos sistemas se pueden definir hasta 11 categorías diferentes que van desde los sistemas basados en reglas, los basados en el conocimiento, las redes neuronales, la metodología orientada a objetos, etc., cada una con sus aplicaciones para diferentes investigaciones y problemas (Liao 2005: 93).

Al intentar aplicar uno de estos sistemas al campo de la arqueología se trata de explorar nuevas vías en la búsqueda de conocimiento, de diseñar estrategias que nos permitan poder hacer un estudio más profundo en menos tiempo, de descubrir información “oculta” en los datos, de hacer “microarqueología”. Se trata de poder disponer de herramientas que ayuden al investigador en su trabajo aportándole ideas e indicios que después habrá de validar. Todo esto no es óbice para que en Arqueología no suceda lo que se está viendo en otros campos, la aparición de sistemas y programas que asisten al investigador en la difícil tarea de la toma de decisiones.

Nos gustaría acabar esta reflexión a futuro con las palabras que nosotros mismos publicamos hace ya varios años:

Con todo, por más que intuyamos la inminencia con que probablemente se producirá el desarrollo de estos sistemas de asistencia experta, debemos ser realistas y reconocer que aún queda mucho trabajo que hacer. Primero está el problema de ponernos de acuerdo dentro de la propia disciplina arqueológica (no podemos automatizar ningún proceso que previamente no tengamos sistematizado desde el punto de vista metodológico) y, en segundo lugar, está la necesidad de abrimos a otros campos como el de la programación informática o el de los sistemas de captura y gestión exhaustiva de datos geométricos (García-Gómez, Fernández de Gorostiza, Mesanza 2011: 28).

## BIBLIOGRAFÍA

- Azkarate Garai-Olaun, A. 2001: “Análisis de la evolución histórico-constructiva de la Catedral de Santa María de Vitoria-Gasteiz (Aplicación de la Arqueología de la Arquitectura a un modelo complejo)” en *V Congreso de Arqueología Medieval Española* (Valladolid, 1999), pp. 177-211. Junta de Castilla y León, Valladolid.
- Azkarate Garai-Olaun, A. 2002: “Intereses cognoscitivos y praxis social en Arqueología de la Arquitectura”, *Arqueología de la Arquitectura*, 1, pp. 55-71. <https://doi.org/10.3989/arq.arqt.2002.6>
- Azkarate Garai-Olaun, A. 2010: “El análisis estratigráfico en la restauración del patrimonio construido. Consideraciones conceptuales e instrumentales”, en E. de Vega y C. C. Martín (ed.), *Arqueología aplicada al estudio e interpretación de edificios históricos. Últimas tendencias metodológicas*, pp. 51-64. Ministerio de Cultura de España, Madrid.
- Azkarate Garai-Olaun, A. 2013: “La construcción y lo construido. Arqueología de la Arquitectura” en J. A. Quirós Castillo (ed.), *La materialidad de la Historia. La arqueología en los inicios del siglo XXI*, pp. 280-308. Akal, Madrid.
- Azkarate Garai-Olaun, A. y Lasagabaster, J. I. 2001: “La catedral de Santa María. Análisis constructivo”, en *I Congreso Europeo de Restauración de Catedrales Góticas (Vitoria-Gasteiz)*, pp. 547-560. Diputación Foral de Álava, Vitoria-Gasteiz.
- Azkarate Garai-Olaun, A., Fernández de Jauregui, A. y Núñez, J. 1995: “Documentación y análisis arquitectónico en el País Vasco. Algunas experiencias llevadas a cabo en Álava-España”, *Informes de Construcción*, 46, 435, pp. 65-77.
- Azkarate Garai-Olaun, A., García-Gómez, I., Domínguez, I. C., Fernández de Gorostiza, M., Gómez, E., Martínez, A., Mesanza, A. y Rodríguez, J. 2010: *Estudio histórico-arqueológico del casco histórico de Vitoria-Gasteiz. Protocolos de intervención, modelos de gestión y Sistema de Información Patrimonial (SWIP)*. Vitoria-Gasteiz.
- Barceló, J. A. 2010: “Computational intelligence in archaeology. State of the art”, en B. Frischer, J. W. Crawford y D. Koller (eds.), *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference*, pp. 11-21. Archaeopress, Williamsburg, Virginia, United States of America. [En línea]: [http://proceedings.caaconference.org/files/2009/02\\_Barcelo\\_CAA2009.pdf](http://proceedings.caaconference.org/files/2009/02_Barcelo_CAA2009.pdf) [consultado el 01/02/2018].
- Caballero Zoreda, L. 2009: “Edificio Histórico y Arqueología: un compromiso entre exigencias, responsabilidad y formación”, *Arqueología de la arquitectura*, 6, pp. 11-19. <https://doi.org/10.3989/arq.arqt.2009.09005>
- Caballero Zoreda, L. 2010: “Experiencia metodológica en Arqueología de la Arquitectura de un grupo de investigación”, en E. de Vega y C. C. Martín (ed.), *Arqueología aplicada al estudio e interpretación de edificios históricos. Últimas tendencias metodológicas*, pp. 103-120. Ministerio de Cultura de España, Madrid.
- Castiñeiras, M. 2012: “San Vicentejo de Treviño, un edificio excepcional en la encrucijada del tardorrománico hispánico”, en R. González de Viñaspre y R. Garay (eds.), *Viaje a Íbiza. Estudios Históricos del Condado de Treviño*, pp. 230-328. Vitoria-Gasteiz.
- García-Gómez, I. et al. 2011: *Montes de Vitoria: Análisis clúster aplicado al estudio estratigráfico de paisajes antropizados*. Vitoria-Gasteiz.
- García-Gómez, I., Fernández de Gorostiza, M. y Mesanza, A. 2011: “Láser escáner y nubes de puntos. Un horizonte aplicado al análisis arqueológico de edificios”, *Arqueología de la Arquitectura*, 8, pp. 25-44. <https://doi.org/10.3989/arq.arqt.2011.10019>
- Hodder, I. y Orton, C. 1990: *Análisis espacial en arqueología*. Crítica, Barcelona.
- Iñiguez Almech, F. 1968: “Sobre tallas románicas del s. XII”, *Príncipe de Viana*, 112-113, pp. 181-235.
- Liao, S. H. 2005: “Expert system methodologies and applications- a decade review from 1995 to 2004”, *Expert systems with applications*, 28, 1, pp. 93-103. <https://doi.org/10.1016/j.eswa.2004.08.003>
- Mannoni, T. y Poleggi, E. 1974: “Fonti scritte e strutture medievali del ‘Castello’ di Genova”, *Archeologia Medievale*, 1, pp. 171-194.
- Mesanza, A. 2017: *Análisis multivariante aplicado a la lectura estratigráfica de alzados*. Tesis doctoral. Escuela Politécnica Superior de Ávila-Universidad de Salamanca.
- Osés, N. y Azkarate, A. 2013: “New protocols for built heritage protection in the Basque Country: towards an automatic analysis tool for built heritage”, en M. A. Rogerio-Candelera, M. Lazzari y E. Cano (eds.), *Science and Technology for the Conservation of Cultural Heritage*, pp. 419-424. CRC Press, Santiago de Compostela, España.
- Osés, N. y Dornaika, F. 2013: “Image-based delineation of built heritage masonry for automatic classification”, en M. Kamel y A. Campilho (eds.), *Image Analysis and Recognition, ICIAR 2013. Lectures in Computer Science*, vol.

- 7950, pp. 782-789. Springer, Berlín, Heidelberg. [https://doi.org/10.1007/978-3-642-39094-4\\_90](https://doi.org/10.1007/978-3-642-39094-4_90)
- Osés, N., Dornaika, F. y Moujahid, A. 2014: "Image-based delineation and classification of built heritage masonry", *Remote Sensing*, 6 (3), pp. 1863-1889. <https://doi.org/10.3390/rs6031863>
- Pérez López, C. y Santín González, D. 2007: *Minería de datos. Técnicas y herramientas*. Ediciones Paraninfo, Madrid.
- Plata, A. 2007: "Nuevas formas de afrontar el estudio del patrimonio salinero: La arqueología de la arquitectura y las salinas de Añana (País Vasco)", en N. Morère Molinero (ed.), *Las salinas y la sal de interior en la historia: economía, medio ambiente y sociedad*, pp. 995-1019. Universidad Rey Juan Carlos, Madrid.
- Portilla, M. y Eguía, J. 1968: *Catálogo monumental de la diócesis de Vitoria. II (Arciprestazgos de Treviño, Albaina y Campezo)*. Obra cultural de la Caja de Ahorros de Vitoria, Vitoria-Gasteiz.
- Rolón, G. 2013: *La vivienda popular riojana del ámbito rural: Patrones arquitectónicos y contexto social en los valles durante el Periodo republicano*. Tesis doctoral. Facultad de Filosofía y Letras-Universidad de Buenos Aires.
- Rolón, G. 2014: "Patrones arquitectónicos, clústeres constructivos homogéneos y variabilidad en el estudio de edificios históricos. Aspectos técnico-formales de la vivienda rural en la provincia de La Rioja (Argentina) durante el periodo republicano", *Arqueología de la Arquitectura*, 11, pp. 1-24. <https://doi.org/10.3989/arq.arqt.2014.013>
- Sánchez Zufiaurre, L. 2007: *Técnicas constructivas medievales. Nuevos documentos arqueológicos para el estudio de la Alta Edad Media en Álava*. Servicio editorial de la Universidad del País Vasco, Leioa.
- Yan, M. 2005: *Methods of determining the number of clusters in a data set and new clustering criterion*. Tesis doctoral. Instituto Politécnico de Virginia y Universidad Estatal. [En línea] <http://hdl.handle.net/10919/29957>